



REC'D 2 5 MAR 1998 WIPO PCT

Patent Office Canberra

I, KIM MARSHALL, MANAGER EXAMINATION SUPPORT AND SALES, hereby certify that the annexed are true copies of the Provisional specification as filed on 20 February 1997 in connection with Application No. PO 5218 for a patent by TELSTRA R & D MANAGEMENT PTY LTD.

I further certify that the annexed documents are not, as yet, open to public inspection.

PRIORITY DOCUMENT

ATENT OFFICE

WITNESS my hand this Second day of March 1998

KIM MARSHALL

MANAGER EXAMINATION SUPPORT AND

SALES

AUSTRALIAN
PROVISIONAL NO. CONTRICT SILING
PO5218 2U FEB. 97
PATENT OFFICE

Telstra R & D Management Pty Ltd

AUSTRALIA
Patents Act 1990

PROVISIONAL SPECIFICATION

for the invention entitled:

"Invisible digital watermarks"

The invention is described in the following statement:

INVISIBLE DIGITAL WATERMARKS

This invention relates to the provision of identification or authentication data, sometimes referred to as a watermark or signature, in digital media data such as digital image or audio data. In particular, the present invention relates to a method and apparatus for incorporating a watermark in digital media data, and a method and apparatus for retrieving or extracting a watermark from digital media data in which a watermark has been previously incorporated.

In this specification the term "watermark" is used to refer to any distinctive or distinguishing

data which may be used for identification or authentication of the digital media data associated therewith, or of some attribute of the media data such as the source thereof. A watermark may comprise image data, such as pixel data forming a logo or the like, or may be in the form of coded text and/or binary numbers, for example, which represent a message. In some applications the watermark data may include error correction coding techniques to improve the robustness of the watermark to image manipulation. The format of the signal that is to be watermarked is not restricted to a multi dimensional representation. It is also possible for audio information to be watermarked. This method of encoding data is not restricted to information associated with copyright and could be used to convey any suitable information in a hidden manner.

20

Watermarks are utilised in media data for a number of reasons, one being to prevent or discourage copying of the media data if it is subject to copyright, or to at least allow for identification of the media data even if it is copied. Visible watermarks have been employed for many years in varying applications including banknotes and photographs, but have significant disadvantages because of their visible nature. Although a visible watermark may be quite effective in discouraging copying of an associated image, in general it is considered disadvantageous for a watermark to be obtrusive upon the original image.

Besides the issue of whether or not the watermark is visible in an associated image (or audible in the case of watermarked audio media), several other factors are also considered important.

For one, the watermark should be robust to manipulation of the watermarked media, and should be secure so as to not be easily removable by a malicious user. Before the advent of digital media processing and manipulation, a degree of robustness and security was inherent in a visible watermark, because a copy of the watermarked image would generally bring with it the visible watermark itself which would be difficult to remove. However, digital processing makes it possible to perform many sophisticated manipulative operations on watermarked media, which may degrade the visible watermark or be utilised to alter an image to at least substantially remove the watermark. In this case, therefore, the properties of a visible watermark count against the security thereof since it is clearly visible what must be removed or altered in the watermarked image. A paper entitled "Protecting publicly—available images with a visible image watermark" (Gordon Braudaway, Karen Magerlein & Fred Mintzer; SPIE Vol. 2659, pp 126–133) discusses robustness and security in visible image watermarks.

Visible watermarks are considered unsuitable for many modern applications because of the intrusive effect of the watermark on the original media. Watermarking schemes have been developed in which the watermark is substantially invisible on an original image but readily visible on a copy thereof. However, such schemes generally rely upon characteristics of photocopying or electronic scanning apparatus, and so are only suitable for a limited range of applications, such as in images or text on paper documents. In any event, these watermarking schemes are also subject to security difficulties arising from digital processing and manipulation.

In media involving a sequence of images, such as video media, it is particularly undesirable for a watermark to be intrusively visible, since considerable effort is expended in providing the image data to the user in a form which is as visually clear as possible, and a visible watermark may significantly detract from the original image. Visible watermarks are presently used in some video applications, particularly television coverage of live sporting events where a relatively small and faint logo or the like is superimposed on the television picture, typically near one corner thereof. This is not completely satisfactory, besides the visual intrusion, because the logo can be easily cropped from the picture in a copy thereof, or

could be relatively easily removed, at least substantially, with digital processing techniques. To make the visible watermark more secure it should be placed over the visually most important part of the image, which also makes the watermark more intrusive and thus less desirable.

5

Invisible watermarking techniques, particularly for digital media data, have been developed, and one is described in an article entitled "Watermarking Digital Images for Copyright Protection" (J.J.K. O'Ruanaidh, F.M. Boland & O. Sinnen). This article discloses a method of embedding a watermark in a digital image which is said to be invisible and quite robust.

The image data is divided into rectangular blocks, and each block is then transformed using a Walsh transform, discrete cosine transform (DCT) or wavelet transform. The bits defining the watermark graphic are inserted in the digital image by incrementing or decrementing a selected coefficient in the transform domain of the data block. Coefficients are selected according to a criterion based on energy content. Another algorithm described in the article relates to insertion of watermark data based on the use of the discrete Fourier transform (DFT). This method differs fundamentally from the transform domain technique outlined above. The DFT is a complex transform that generates complex transform domain coefficients given a real valued input. The watermark is placed in the phase component of generated transform coefficients when using this transform.

20

Another article which addresses the difficult issues of digital watermarking is "Secure Spread Spectrum Watermarking for Multimedia" (Ingemar J Cox, Joe Kilian, Tom Leighton & Talal Shamoon; NEC Research Institute, Technical Report 95-10). This article describes an invisible digital watermarking method for use in audio, image, video and multimedia data.

25 The method described in this article also involves a frequency domain transform of the image data and insertion of the watermark data whilst in the transform domain. In practice, in order to place a length n watermark into an N x N image, the discrete cosine transform of the image is computed, and the watermark data encoded into the n highest magnitude coefficients of the transform matrix, excluding the dc component. In other words, the watermark data is placed in transform domain components of greatest perceptual significance, which enables the

watermark to be robust to image distortion and unauthorised removal without serious degradation of the image itself. This watermarking algorithm employs an energy compacting transform, which makes the selection of transform coefficients for encoding of the watermark data very important. For most images the coefficients selected will be the ones corresponding to the low spatial frequencies, with the result that significant tampering of the image at those frequencies would destroy the image fidelity before the encoded watermark. The watermarking techniques of J.J.K O'Ruanaidh et al and Ingemar J. Cox et al require the original image when performing the watermark extraction operation. As a consequence, proof of ownership is accomplished only if the original image is certified as being the original by a trusted third party, and the particular segment of the original image must be first identified and found before ownership is verified.

The present invention addresses some of the difficulties identified in the prior art, and embodiments of the invention aim to provide a digital watermarking process in which:

15

- 1. the presence of the watermark is invisible (i.e. the watermarked visual or audio material is visually or auditorially substantially indistinguishable from the original);
- 2. the watermark is robust to signal manipulation;

20

- 3. the watermark is secure;
- 4. the original media data is not required in order to extract the watermark; and
- 25 5. the watermark can be inserted and/or extracted by a simple computational procedure which can be done in real time on the fly.

In accordance with the present invention, there is provided a method for inserting identification or authentication data into digital media data, comprising the steps of:

segmenting the digital media data into data blocks;

permuting a block of the digital media data;

performing a transform on the permuted block of digital media data to obtain transform domain data;

modifying at least one component of the transform domain data to incorporate 5 identification/authentication data;

inverse transforming the transform domain data with modified component; and performing an inverse permute to obtain watermarked digital media data.

Preferably, each coefficient generated by the transform contributes substantially equally to the total energy of the block. As a consequence, the watermarking process is not sensitive, with regard to the introduced distortion, to the selection of the transform coefficient which is modified in the watermark insertion operation.

15 Preferably the insertion method is performed in real time with respect to the digital media data.

The present invention also provides a method for extracting identification or authentication data from watermarked digital media data, comprising the steps of:

segmenting the digital media data into data blocks;

permuting a block of the digital media data;

performing a transform on the permuted block of digital media data to obtain transform domain data; and

extracting a component of the transform domain data to obtain 25 identification/authentication data.

The present invention also provides an apparatus for inserting or extracting watermark data in digital media data, comprising:

segmenting means for segmenting the digital media data into data blocks;

processing means for permuting a block of the digital media data and performing a



30

transform on the permuted block of digital media data to obtain transform domain data; and means for inserting or extracting watermark data in a component of the transform domain data.

- 5 Preferably, in the case where watermark data is to be inserted in the digital media data the processing means is also adapted to perform an inverse transformation and inverse permute of the transform domain data containing the watermark data so as to obtain watermarked digital media data.
- 10 In practice, the segmenting of the digital media data into data blocks might comprise of forming blocks of 64x64 pixels of image luminance pixel data, where the watermark is to be inserted into a still image or image sequence. The block size need not be restricted to being square and of dimension 64x64 pixels, both smaller and larger block sizes are possible depending upon application requirements. In practice, the identification/authentication data which is inserted into a data block of digital media data might comprise a pixel from a binary graphic, or data in the form of bits used to represent text and binary numbers, for example. The watermark data is inserted into the data block that has undergone a block transform operation. The distortion introduced due to the insertion of watermark data is dependent upon the block size, the number of transform coefficients modified by the insertion operation and the magnitude of the modification. The watermark data density per block is arbitrary depending upon application requirements. In general, however, the higher the density the more visually noticeable is the inserted watermark in the image. A series of data blocks may be contained in a single image frame or spread over a number of image frames.
- 25 The invention is described in greater detail hereinafter, by way of example only, with reference to the accompanying drawings, wherein:

Figure 1 is a flowchart illustrating operations for inserting watermark data into digital media data;

Figure 2 is a flowchart illustrating operations for extracting watermark data from 30 digital media data;

5

20

Figure 3 is a diagram of the watermark insertion process of a preferred embodiment of the present invention;

Figure 4 is a flowchart illustrating the operations for a particular implementation of the watermarking insertion procedure; and

Figure 5 is a block diagram of watermarking apparatus for real-time video.

This invention relates to the insertion and extraction of identification or authentication data for use as a watermark in digital media data, such as digital image data, still or sequential, digital audio data or the like. A watermark provided in digital media data may provide a means for identification of the source or some other attribute of the media data as may be required to prove copyright ownership, for example. As mentioned above, embodiments of the present invention are designed to have a number of advantageous properties, including:

the presence of the watermark being at least substantially invisible (ie the watermarked visual or audio material is visually or auditorially substantially indistinguishable from the original);

the watermark can be inserted and/or extracted by a simple computational procedure which can be done in real time; and

the original media data not being required in order to extract the watermark from the watermarked media data.

Additionally, as also discussed above, it is desirable for watermarks in digital media data to also be both secure in that a malicious user cannot easily remove or disguise the watermark so as to prevent extraction, and robust to enable the inserted watermark to survive manipulation of the watermarked media data. Digital images and image sequences, for example, are seldom stored or transported over a communications link their raw format. Frequently some form of compression may be applied to the media data, and it is therefore important that the signal processing associated with the compression algorithm does not remove or wash out the associated watermark inserted in the media data.

30 Although the following description of embodiments of the present invention refer primarily

to still or sequential image data, it is to be understood that the invention is equally applicable to other forms of digital media data, such as digitised audio data.

In an embodiment of the invention, image pixel data is subdivided into 64 x 64 pixel spatial domain blocks in order to provide a manageable data segment in which to insert watermark data. For example, a digital image comprising 1,024 x 768 pixels may be nominally divided into blocks of 64 x 64 pixels so that the entire image is contained in an array of 16 x 12 image data blocks (a total of 192 data blocks). Different watermark data may then be inserted into each data block, so that the watermark data is spread over the entire image. For example, the watermark might comprise a 16 x 12 pixel logo or the like, so that a value representing each pixel of the logo is inserted in a respective data block of the digital image. Alternatively, the watermark may comprise a text message formed in ASCII code and/or binary numbers. A message comprising of 192 bits could be inserted in the digital image if a watermark density of 1/4096 (one bit per 64x64 block) was employed.

15

The invisibility and robustness of the watermark are aided by dividing the image into blocks and distributing the watermark data throughout the data blocks, and are further facilitated by the insertion procedure utilised to insert the watermark data into each data block. The following steps are used to insert a watermark data bit or binary pixel graphic into a 64 x 64 20 spatial domain data block.

- (I) The average pixel value for the 64x64 data block is computed and subtracted from each pixel value.
- 25 (ii) The 64x64 data block is then transformed using a variant of the Walsh Hadamard Transform referred to in the following description as the Energy Spreading Transform (EST). The EST of order 4 is given by:

5

1.

An EST of higher order is determined by a recursive rule for values of N that are an integer multiple of 4. A 64 point transform is obtained with N = 16.

2

- This transform is characterised by having energy spreading properties, rather than energy compaction properties associated with classic block transforms. Natural images in general have a low pass power spectral density, and thus the energy spreading transform is utilised so as to generate a block of transform coefficients that have comparable magnitude values.

The transform operation is two dimensional and separable.

(iii) The watermark data is inserted into the data block by modification of the selected transform coefficient/s. As each transform coefficient contributes approximately an equal

15

energy contribution to the data block, the selection process can be a deterministic pseudo random procedure. The distortion introduced by this watermarking procedure is not sensitive to which transform coefficient is selected and modified.

- 5 A watermark data bit could be represented by the sign of a selected transform coefficient. A transform coefficient value greater than or equal to zero could represent logic zero and the negative values logic one. Transform coefficient/s need only be modified if necessary, to ensure that the sign (+/-) corresponds the digital bit to be embedded (1/0).
- 10 (iv) An inverse transform is then applied and the previously subtracted average pixel valued added to each pixel to reconstruct an approximation of the original 64 x 64 spatial domain data block. In the transform domain, the watermark data is contained completely by one transform coefficient when using a watermark data density 1/4096. In the spatial domain, however, the watermark data is distributed over each of the pixels making up the 64x64 data block.

The watermark read operation is accomplished by repeating steps (I) and (ii) above. The watermark data can be extracted with the knowledge of only the transform and the transform coefficient/s that have been modified to contain the watermark data. The original image or image sequence is not required for the reading operation. This watermarking technique has demonstrated a high degree of robustness to coding artefacts. Useful watermarks have been extracted using this approach following distribution quality MPEG 2 encoding that employs forward and backwards prediction. However, the security of the watermarking process described in detail hereinabove is not particularly strong when utilised in watermarking image sequences. There exists a high degree of temporal correlation between frames of an image sequence. If the transform is known, it is possible to identify transform coefficients that contain watermark data thus allowing the watermark to be removed or made void.

A much more secure watermarking procedure is obtained by first permuting the spatial domain data prior to the transform operation. For a 64 x 64 data block there exists 4096 factorial different ways in which the permutation can be performed. To extract and read or remove the

watermark requires knowledge of the permutation employed which is preferably kept secret by the owner of the image or image sequence. This could be in the form of a secret seed number to a well defined and widely known pseudo random number generator. The security of this watermarking procedure is, as a consequence significantly enhanced. A secondary 5 feature of the permutation operation is to flatten the power spectral density of the data block. Accordingly, the advantageous energy spreading properties discussed above need not necessarily be provided by the transform operation, which enables other transforms to be utilised in place of the EST. Block transforms such as the classic Walsh Hadamard Transform (WHT), Discrete Cosine Transform (DCT), Discrete Sine Transform (DST) and the Haar 10 Transform (HT) can therefore be employed. For transforms that isolate the average block value into one coefficient it is required that this coefficient is not used to contain watermark data. The WHT or the EST are, however, the preferred choice for the transform operation due to their low implementation complexity. Fast transform implementations of the WHT and EST exist that require only summing and one scaling operation, as both transforms are 15 characterised by basis vectors which contain only +1 and -1 elements. The analysis and synthesis transforms are identical. The WHT contains the average value of the block in the DC coefficient, and hence this coefficient cannot be used to contain watermark data. As a consequence, however, the block average does not need to be subtracted from each data element of the media data segment. This is not the case with the EST, and hence the average 20 subtraction operation should be performed for optimal results, but all of the transform coefficients are of course available for modifiaction to contain watermark data. Both transforms produce a robust, secure and substantially invisible watermark using the procedure which is detailed herein.

25 Figure 1 illustrates a flow chart of operation involved in insertion of watermark data into digital media data, according to an embodiment of the invention. Beginning at step 12, the digital media data is first segmented into manageable data blocks such as blocks of 64x64 pixels or equivalent data elements. Step 13 calculates the average pixel value for the block which is then subtracted from each pixel. Step 13 is unnecessary when using a transform that 30 contains the block average in a single transform coefficient. This is the case with the WHT

and the DCT. The resulting DC transform coefficient cannot however be used to contain watermark data. The media data block or segment is then subjected to a permute operation (step 14) in which the data elements of the block or segment are rearranged in a pseudo random, but repeatable and reversible manner. Next, at step 16, the permuted spatial domain 5 media data segment is subjected to the transform operation. In this embodiment one of the transform coefficients is selected and modified to include watermark data. When watermarking images or image sequences a watermark data bit could be represented by the sign of the selected transform coefficient. A transform coefficient value greater than or equal to zero could represent logic zero and the negative values logic one.

10

The watermark data density per block in this case is 1/4096. In some applications, densities greater than 1/4096 may be required.

Following insertion of the watermark data into the transform domain of the media data, the spatial domain media data is then reconstructed through steps 20, 22 and 23 by performing an inverse transformation followed by an inverse permute operation and then the previously subtracted block average value added to each pixel making up the block. Again, step 23 is not necessary when using a transform that contains the block average in a single transform coefficient. The resulting digital media data segment contains watermark data which is robust to manipulation thereof, secure from unauthorised removal, and yet the reconstructed, watermarked media data is substantially indistinguishable from the original spatial domain media data when compared in subjective quality testing.

In order to extract the watermark data form digital media data in which watermark data has been previously inserted, the procedure outlined in the flow chart of Figure 2 may be employed. Essentially this involves steps mirroring the first half of the procedure illustrated in Figure 1. The digital media data is first segmented as discussed previously (step 32), the average pixel value for that block is determined and subtracted from each pixel (step 33) if necessary. The resulting data block is then subjected to a permute operation as shown at step 34. The permute operation must be the same as that performed during insertion of the

watermark data, and thus if different permute operations are variously employed, some record must be maintained of which of the particular 4096 factorial permutations applies to the particular media data segment in question. This could be in the form of a secret seed to a well defined and widely known pseudo random number generator. The permuted media data segment is then transformed with the same transform used by the insertion operation (step 36). Then it is a simple matter to extract the particular coefficient for the transform domain media data and then recover from this the watermark information.

Figure 3 illustrates a block diagram of the watermark insertion process described in connection with the flow chart of Figure 1. As discussed above, in this embodiment only a single watermark data component, eg a data bit or binary graphic pixel, is inserted into each selected digital media data segment or block, and the information required to reconstruct an entire watermark requires the examination of a number of digital media data segments.

. 3

15 Figure 4 is a flow chart illustrating the insertion process of watermark data into digital media data, which has been segmented into data blocks, over a series of data blocks. Where the digital media data comprises a sequence of images, such as in the case of digital video or the like, a complete watermark (eg the total of the identification data) may in fact be distributed over more than one image or image frame. At step 42 the first data block in the image or 20 sequence of images is selected and, if necessary, the average of that block is then calculated and subtracted from each pixel element in step 43. The resulting data block forming the image segment is subjected to a permute operation, as described hereinabove, at step 44. The permuted image data is then transformed using a block transform. At step 47 a particular transform coefficient is selected for possible modification. The selection process is performed 25 in a pseudo random deterministic manner. Transforms that contain the block average in one transform coefficient must eliminate this coefficient from the selection process. Step 48 performs the modification operation to incorporate the watermark data into the selected transform coefficient. The inverse of the transformation and permute operations are then applied at steps 50 and 52 and step 53 adds to each pixel value the average as determined in 30 step 43, if necessary. A test is then applied at step 54 to determine whether the media data has

finished, and if so the watermarking procedure ends. Otherwise, the next block of the digital media data is selected at step 56. The watermark data is then incremented, meaning the next component of the watermark data, such as the next data bit or binary pixel element, is selected at step 58. Of course, it will be recognised that it is unnecessary for every data block of a particular digital media data source to be encoded with watermark data, and only a certain selection of data blocks may in fact be encoded with watermark data in practice. To provide copyright protection for the complete image sequence, the watermark can be repeatedly inserted, with the watermark beginning at different frame locations within the sequence and ensuring that watermarks do not overlap. Of course, acquisition of the signal is important.

This can be accomplished, by incorporating in the watermark data, synchronisation information that, once acquired informs the watermark reader the location of the beginning of the watermark message data or binary graphic.

To increase robustness and ensure readability even in the case where the original video signal is significantly changed, such as through reduced spatial resolution or the case where watermarked interlaced material is later converted to non-interlaced format, the watermark can be distributed across both fields in such a way that the watermark can be independently read from either or both fields and/or restricted to the low spatial frequencies. The latter can be accomplished by the application of a 2x2 WHT on each row of the image to produce low and high spatial frequency components. The watermark is then inserted in only the half horizontal resolution frame corresponding to the low spatial frequencies. The full resolution watermarked frame is produced by performing an inverse 2x2 WHT on the rows making up the low spatial frequency watermarked half horizontal resolution frame and the original high spatial frequency half horizontal resolution frame.

25

In order to further improve security of the watermarking procedure, it is possible to alter the permute operation periodically. As mentioned above, it is nevertheless necessary that the particular permute operation performed on each data block be repeatable at a future time to enable extraction of the watermark.

Figure 5 illustrates a block diagram of watermarking apparatus for encoding real time video with watermark data according to an embodiment of the present invention. Real time video feed is provided to the apparatus at a buffer 80 or the like, which provides an input to real time processing circuitry 82. The circuitry 82 may comprise digital processing circuitry in the form of high speed programmable computer circuitry, for example, which carries out the algorithmic steps illustrated in Figure 4. The watermark data is provided from a buffer 84 which may be in the form, for example, of a ring buffer which cyclically feeds watermark data being a component of watermark text or graphic material to the processing circuitry 82. The reconstructed video data containing the watermark data is then passed to an output buffer 86 which provides the video data for transmission, recording or whatever function the video data is required for.

It will be appreciated from the foregoing description that the original media data is not required by the watermark extraction process in order to extract the watermark data, and 15 therefore it is not required that the original image be certified by a trusted third party or held in escrow in order to prove the presence of a watermark in the media data. Random accessibility of a watermark within an image sequence is easily achieved, as all that is required to extract the watermark is the image or sequence of images that contains sufficient watermark data to reconstruct the entire watermark or a substantial portion thereof. The 20 watermarking process according to an embodiment of the invention has been tested on still images and image sequences, and has been demonstrated to be near invisible to the naked eye in a comparison between the reconstructed, watermarked media data and the original media data. It has also be found to be secure and robust to compression such as 4 Mbps MPEG coding of image sequences and 20% quality setting for JPEG compressed still images. It will 25 also be appreciated that the simple nature of the computational processes involved in the watermarking process of the present invention allow it to be applied quite readily to real time video data, for example. This is because the only two computationally complex steps in the watermarking procedure, namely the permute and transformation are still relatively simple. This makes for a watermarking process that is very low in complexity, is easily automated, 30 and requires no human intervention in its application.

The foregoing detailed description of the present invention has been presented by way of example only, and is not intended to be considered limiting to the invention, which includes every novel feature and novel combination of features herein disclosed.

5

Dated this 20th day of February 1997

10 TELSTRA R & D MANAGEMENT PTY LTD

By its Patent Attorneys

DAVIES COLLISON CAVE

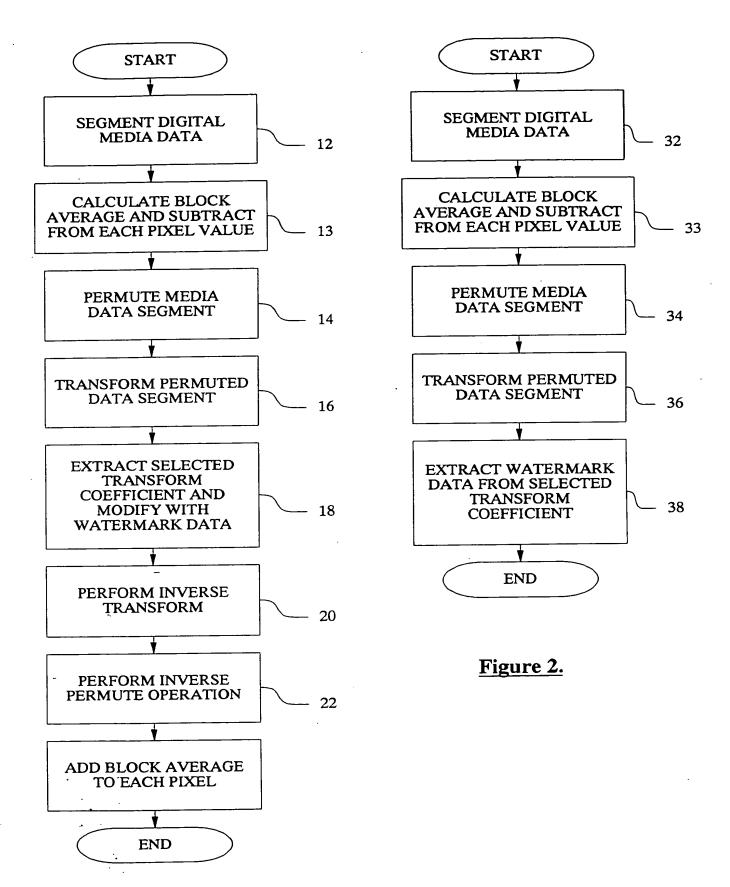


Figure 1.

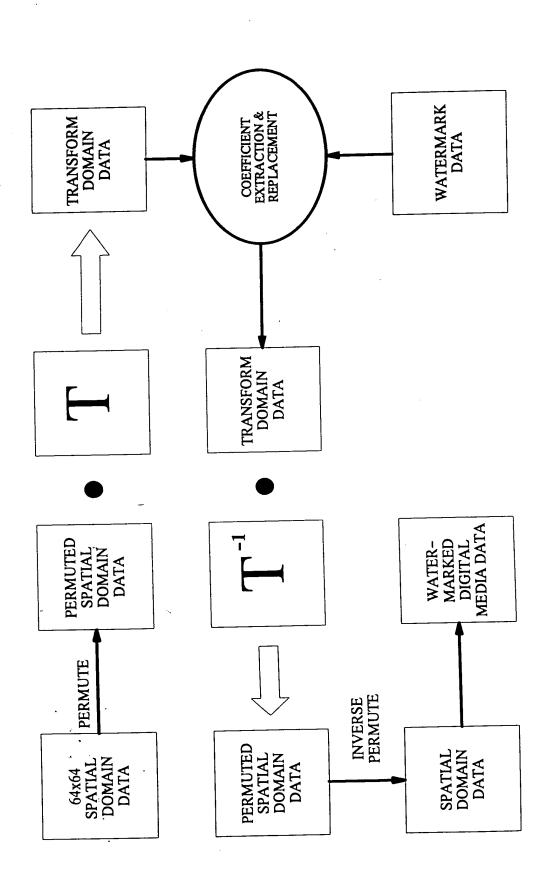


Figure 3.

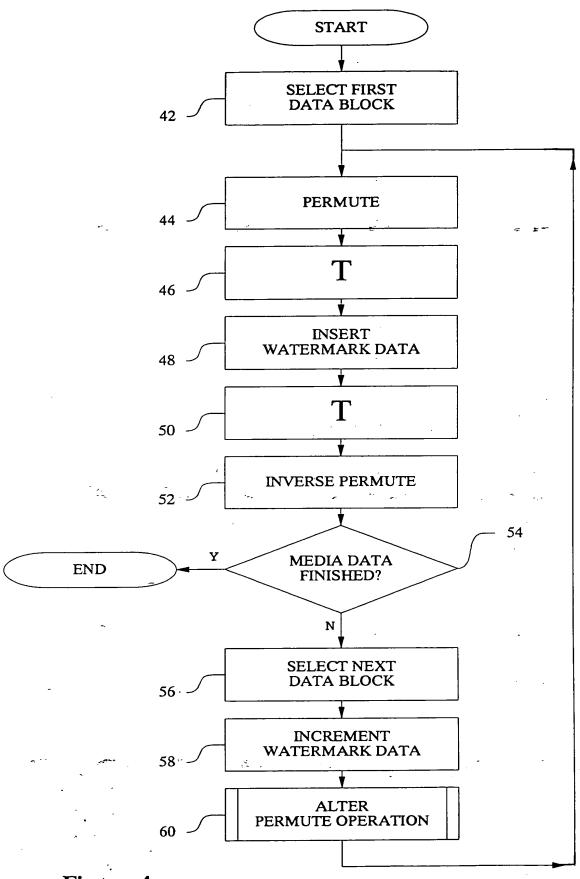


Figure 4.

. **L**.

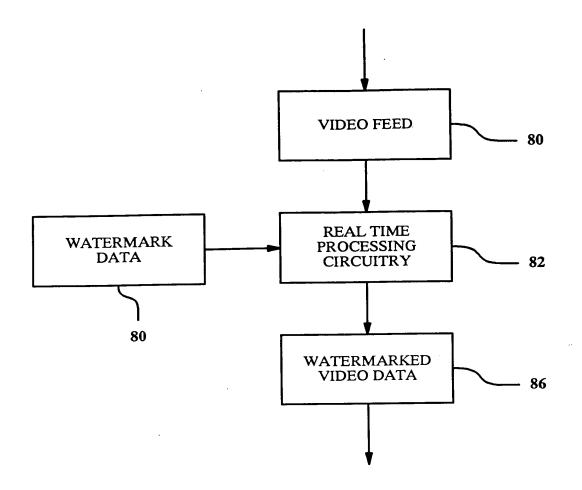


Figure 5.